Rectangular R-transform as the limit of rectangular spherical integrals

Florent Benaych-Georges

RECTANGULAR R-TRANSFORM AS THE LIMIT OF RECTANGULAR SPHERICAL INTEGRALS

FLORENT BENAYCH-GEORGES

Abstract. In this paper, we connect rectangular free probability theory and spherical integrals. We prove the analogue, for rectangular or square non-Hermitian matrices, of a result that Guionnet and Maïda proved for Hermitian matrices in [12]. More specifically, we study the limit, as n, m tend to infinity, of $\frac{1}{n}\log \mathbb{E}\{\exp[\sqrt{nm}\theta X_n]\}$, where $\theta \in \mathbb{R}$, X_n is the real part of an entry of $U_n M_n V_m$, M_n is a certain $n \times m$ deterministic matrix and U_n, V_m are independent Haar-distributed orthogonal or unitary matrices with respective sizes $n \times n$, $m \times m$. We prove that when the singular law of M_n converges to a probability measure μ, for θ small enough, this limit actually exists and can be expressed with the rectangular R-transform of μ. This gives an interpretation of this transform, which linearizes the rectangular free convolution, as the limit of a sequence of log-Laplace transforms.

Introduction

In this article, we study the limit, as n, m tend to infinity in such a way that n/m tends to a limit $\lambda \in [0, 1]$, of

$$\frac{1}{n}\log \mathbb{E}\{\exp[\sqrt{nm}\theta\, \Re(\mathrm{Tr}(E_n U_n M_n V_m))]\},$$

where $\theta \in \mathbb{R}$, M_n is a certain $n \times m$ deterministic matrix, U_n, V_m are independent Haar-distributed orthogonal or unitary matrices with respective sizes $n \times n$, $m \times m$, E_n is an $m \times n$ elementary matrix (i.e. a matrix which entries are all zero, except one of them, which is equal to one) and $\Re(\cdot)$ denotes the real part.

The departure point of this study is the work of Collins, Zinn-Justin, Zuber, Guionnet, Maïda, Śniady, Mingo and Speicher who proved, in the papers [20, 6, 12, 7, 8, 9], that under various hypotheses on some $n \times n$ matrices A_n and B_n and a positive exponent α, the asymptotics of

$$\frac{1}{n^\alpha}\log \mathbb{E}\{\exp[n\theta\, \mathrm{Tr}(B_n U_n A_n U_n^*)]\}$$

are related to free probability theory. For example, it has been proved [12, Th. 2] that if the spectral law (i.e. uniform distributions on eigenvalues) of the self-adjoint matrix

Date: April 19, 2011.

2000 *Mathematics Subject Classification.* 15A52, 46L54, 60B15.

Key words and phrases. Random Matrices, Free Probability, Rectangular R-transform, Haar Measure, Spherical Integrals.

This work was partially supported by the *Agence Nationale de la Recherche* grant ANR-08-BLAN-0311-03.

A_n converges to a compactly supported probability measure μ and $F_n = \operatorname{diag}(1, 0, \ldots, 0)$, then for θ small enough,

$$\frac{1}{n} \log \mathbb{E}\{\exp[n\theta \operatorname{Tr}(F_n U_n A_n U_n^*)]\} \underset{n\to\infty}{\longrightarrow} \frac{\beta}{2} \int_0^{\frac{2\theta}{\beta}} R_\mu(t)\mathrm{d}t, \tag{1}$$

where R_μ is the so-called *R-transform* of μ and $\beta = 1$ or 2 according to wether we consider real or complex matrices. The R-transform is an integral transform of probability measures on $\mathbb{R}$. Its main property is that it linearizes the additive free convolution $\boxplus$, the binary operation on probability measures on $\mathbb{R}$ which can be defined by the fact that for A, B large self-adjoint random matrices with spectral laws tending to μ_A, μ_B (as the dimension goes to infinity) and U a Haar-distributed orthogonal or unitary matrix independent of A and B, the spectral law of $A + UBU^*$ tends to $\mu_A \boxplus \mu_B$: the free convolution $\boxplus$ can be thought as the analogue, for the spectral laws of certain large random matrices, of the classical convolution for real random variables. For all probability measures μ, ν on $\mathbb{R}$, we have

$$R_{\mu\boxplus\nu}(t) = R_\mu(t) + R_\nu(t) \qquad \text{(for } t \text{ in a neighborhood of zero)}. \tag{2}$$

Hence in the analogy between the free convolution and the classical one, the R-transform plays the role of the log-Laplace transform, and (1) gives a concrete sense to this: the R-transform (more specifically its primitive, which also satisfies (2)), is the limit of a certain sequence of log-Laplace transforms.

Let us now describe the content of our paper. For each $\lambda \in [0, 1]$, another free convolution, denoted by $\boxplus_\lambda$ and called the *rectangular free convolution with ratio* λ, does the same job as $\boxplus$ for the *singular laws* (i.e. uniform distributions on singular values) of rectangular $n \times m$ random matrices which dimensions n, m tend to infinity in such a way that n/m tends to λ: for n, m large integers such that $n/m \simeq \lambda$, for A, B some $n \times m$ real or complex matrices with singular laws ν_A, ν_B and U, V Haar-distributed orthogonal or unitary matrices independent of A and B, the singular law of $A + UBV$ is approximately $\nu_A \boxplus_\lambda \nu_B$ (see [4] or the introduction of [5] for a more precise definition of $\boxplus_\lambda$). Like the R-transform for $\boxplus$ and the log-Laplace transform for the classical convolution, an integral transform linearizes $\boxplus_\lambda$. It is called *the rectangular R-transform with ratio* λ and is denoted by $C^{(\lambda)}$: for all probability measures μ, ν on $[0, +\infty)$, we have

$$C^{(\lambda)}_{\mu\boxplus_\lambda\nu}(t) = C^{(\lambda)}_\mu(t) + C^{(\lambda)}_\nu(t) \qquad \text{(for } t \text{ in a neighborhood of zero)}. \tag{3}$$

The main result of the paper gives an interpretation of the rectangular R-transform (more specifically its primitive, which also satisfies (3)) as the limit of a sequence of log-Laplace transforms: we prove that if the singular law of M_n tends to a probability measure μ and n/m tends to a limit $\lambda \in [0, 1]$ as n, m tend to infinity, then for θ small enough, for E_n a sequence of $m_n \times n$ elementary matrices,

$$\frac{1}{n} \log \mathbb{E}\{\exp[\sqrt{nm}\theta\, \Re(\operatorname{Tr}(E_n U_n M_n V_m))]\} \underset{n,m\to\infty}{\longrightarrow} \beta \int_0^{\frac{\theta}{\beta}} \frac{C^{(\lambda)}_\mu(t^2)}{t}\mathrm{d}t. \tag{4}$$

Let us mention that free probability theory has initially been built in the area of operator algebras and that concrete relations between *free* and *classical* probability theory, like the ones of (1) and (4), are not that common.

Let us also mention that expectations of the exponential of traces of polynomials of constant matrices and uniform orthogonal random matrices, which have been extensively studied in physics and also other areas, like information theory, are often called *spherical integrals*. See e.g. [21, 11, 13] and the references above for the case of square matrices. Spherical integrals involving rectangular matrices are considered in the papers [19, 10, 16]. The quantities studied in the paper [16] of Kabashima are closely related to the spherical integral we study here: in Equation (8), Kabashima considers exactly the same spherical integral as ours, with the same hypotheses (he supposes the singular law of X to have a limit and p/N to stay bounded), except that he supposes θ to be on the imaginary line, whereas in our paper, θ is real. He is not giving an argument that qualifies as a mathematical proof, but he gives an asymptotic formula in Equation (9). With our vocabulary, this formula expresses

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}\{\exp i \operatorname{Tr}(E_n U_n M_n V_m)\}$$

as a saddle point value of a certain function. The actual computation of this saddle point is not easy, but his Equation (9) has a structure which is quite close to our Equation (29).

The paper is organized as follows. In Section 1, we state the main result of the paper, Theorem 1.2, and discuss it. In Section 2, we recall the precise definition of the rectangular R-transform and prove a result of continuity of the map $(\lambda, \mu) \longmapsto C_\mu^{(\lambda)}$. At last, Section 3 is devoted to the proof of Theorem 1.2, following the ideas of the proof of [12, Th. 2].

1. Main result

1.1. **Statement.** Let us consider, for all $n \geq 1$, an integer $m_n \geq n$ such that, as n tends to infinity, n/m_n tends to a limit $\lambda \in [0, 1]$ and an $n \times m_n$ real or complex nonrandom matrix M_n whose singular values are strictly bounded, uniformly in n, by a constant K and such that, as n tends to infinity, the singular law of M_n converges weakly to a probability measure that we shall denote by μ. Let us define, for $\theta \in \mathbb{R}$,

$$I_n(\theta) = \frac{1}{n} \log \mathbb{E}\{\exp[\sqrt{nm_n}\theta\, \Re(\operatorname{Tr}(E_n U_n M_n V_n))]\}, \tag{5}$$

where U_n, V_n are independent Haar-distributed orthogonal or unitary (according to whether M_n is real or complex) matrices with respective sizes $n \times n$, $m_n \times m_n$ and E_n denotes an $m_n \times n$ elementary matrix (i.e. a matrix which entries are all zero, except one of them, which is equal to one).

In the case where $\lambda = 0$, we also suppose that there is $\alpha < 2$ such that

$$\text{for } n \text{ large enough,} \quad m_n \leq n^\alpha. \tag{6}$$

Remark 1.1. Let $\mathbb{K}$ be either $\mathbb{R}$ or $\mathbb{C}$ according to wether we consider real or complex matrices. $I_n(\theta)$ can also be considered as the Laplace transform of a certain scalar product estimated at a pair of independent random vectors, one of them being a uniform random vector of the unit sphere of $\mathbb{K}^n$ and the other one being the projection, on $\mathbb{K}^n$, of a uniform random vector of the unit sphere of $\mathbb{K}^{m_n}$. Indeed, let us denote the singular values of M_n

by $\mu_{n,1},\ldots,\mu_{n,n}$ and introduce (see [15]) some orthogonal or unitary matrices P_n, Q_n with respective sizes $n\times n$, $m_n\times m_n$ such that

$$M_n = P_n \begin{bmatrix} \mu_{n,1} & & & 0 & \cdots & 0 \\ & \ddots & & \vdots & & \vdots \\ & & \mu_{n,n} & 0 & \cdots & 0 \end{bmatrix} Q_n.$$

Let also, for each n, (i_n, j_n) be the index of the non-null entry of E_n. Then the j_nth row (resp. i_nth column) $u_n = (u_{n,1},\ldots,u_{n,n})$ (resp. $v_n = (v_{n,1},\ldots,v_{n,m_n})^t$) of U_nP_n (resp. Q_nV_n) is uniformly distributed on the unit sphere of $\mathbb{K}^n$ (resp. $\mathbb{K}^{m_n}$) and one has

$$I_n(\theta) = \frac{1}{n}\log \mathbb{E}\{\exp[\sqrt{nm_n}\theta\,\Re(\sum_{k=1}^{n} u_{n,k}\mu_{n,k}v_{n,k})]\}. \tag{7}$$

The main result of the article is the following one.

Theorem 1.2. *Set $\beta = 1$ or $\beta = 2$ according to wether we consider real or complex matrices. The function I_n converges uniformly on every compact subset of $(-\beta K^{-1}, \beta K^{-1})$ to the function*

$$I(\theta) = \beta \int_0^{\frac{\theta}{\beta}} \frac{C_\mu^{(\lambda)}(t^2)}{t}\mathrm{d}t,$$

where $C_\mu^{(\lambda)}$ denotes the rectangular R-transform of μ with ratio λ (its definition is recalled in Section 2 below).

Remark 1.3. Note that the function $C_\mu^{(\lambda)}$ is analytic on $(-K^{-2}, K^{-2})$ and vanishes at zero, so I is actually well defined and analytic on $(-K^{-1}, K^{-1})$.

1.2. **Particular cases where the matrices M_n are square ($\lambda = 1$) or asymptotically flat ($\lambda = 0$).** Let us recall that the *R-transform* of a probability measure ν is the function

$$R_\nu(z) = G_\nu^{-1}(z) - \frac{1}{z}, \qquad \text{for } G_\mu(z) = \int \frac{\mathrm{d}\mu(t)}{z-t}$$

(the convention we use is the one of the analytic approach to freeness [14, 1], which is not exactly the one of the combinatorial approach [18]: $R_\nu^{\text{combinatorics}}(z) = zR_\nu^{\text{analysis}}(z)$).

Let μ_s be the symmetrization of μ, defined by $\mu_s(A) = \frac{\mu(A)+\mu(-A)}{2}$ for all Borel subset A of $\mathbb{R}$, and μ^2 be the push-forward of μ by the function $t\mapsto t^2$.

Corollary 1.4. *In the particular case where $\lambda = 1$ (resp. $\lambda = 0$), the limit I of I_n can be expressed via the R-transform of μ_s (resp. μ^2) in the following way*

$$I(\theta) = \beta\int_0^{\frac{\theta}{2}} R_{\mu_s}(t)\mathrm{d}t, \qquad \textit{(resp. } I(\theta) = \beta\int_0^{\frac{\theta}{2}} tR_{\mu^2}(t^2)\mathrm{d}t.\textit{)}$$

Proof. It suffices to prove that $C_\mu^{(1)}(t^2) = tR_{\mu_s}(t)$ and that $C_\mu^{(0)}(t) = tR_{\mu^2}(t)$. The second equation can be found in [4, Lem. 3.2 or Sect. 3.6]. The first equation follows from the fact that for all λ, $C_\mu^{(\lambda)} = C_{\mu_s}^{(\lambda)}$ and from the fact that for all symmetric probability measure ν, by [4, Sect. 3.6], $C_\nu^{(1)}(z^2) = zR_\nu(z)$. $\square$

1.3. **Possible extensions of Theorem 1.2.**

1.3.1. *Cumulants point of view.* For $\lambda \in [0,1]$, the *rectangular free cumulants with ratio* λ of μ have been defined in [4, Sect. 3.4] (see also [2, Sect. 2.2]): this is the sequence $(c_{2k}(\mu))_{k\geq 1}$ linked to the moments of μ by [3, Eq. (4.1)]. Recall also that for X a bounded real random variable, the *classical cumulants* of X are the numbers $\mathrm{Cl}_k(X)$ defined by the formula

$$\log \mathbb{E}(e^{zX}) = \sum_{k\geq 1} \frac{\mathrm{Cl}_k(X)}{k!} z^k.$$

Differentiating formally the convergence $I_n(\theta) \longrightarrow I(\theta)$, one would get the following "classical cumulants interpretation" of the rectangular free cumulants with ratio λ: for all positive integers k,

$$c_{2k}(\mu) = \lim_{n\to\infty} \frac{(nm_n)^k}{n} \frac{\beta^{2k-1}}{(2k-1)!} \mathrm{Cl}_{2k}(\Re(\mathrm{Tr}(E_n U_n M_n V_n))). \tag{8}$$

This formula can be considered as a "rectangular analogue" of [6, Th. 4.7]. The author believes that (8) can be proved rigorously with one of the following two methods. One of them would be to use the *Weingarten calculus*, developed by Collins and Śniady, for the computation of the expectation of moments of the entries of the matrices U_n, V_n (as in the proof of [6, Th. 4.7]). The other one would rely on the extension of Theorem 1.2 to complex values of θ and notice that the functions in question there are analytic in θ (so that their convergence implies the one of their derivatives).

1.3.2. *Case where M is chosen at random.* If M_n is also chosen at random, independently of U_n and V_n, and the expectation, in (5), is taken with respect to the randomness of U_n, V_n and M_n, then Theorem 1.2 stays true in certain cases (for example if M_n is a standard Gaussian matrix divided by $\sqrt{m_n}$), but it can easily be seen that Theorem 1.2 is not true in general anymore. However, if the expectation, in (5), only concerns the randomness of U_n and V_n, then $I_n(\theta)$ is a random variable, and for certain sequences of random matrices M_n, more than Theorem 1.2 can be said about its convergence. An important example is given by the case where $M_n = A_n + P_n B_n Q_n$ with A_n, B_n deterministic matrices having limit singular laws μ_A, μ_B and P_n, Q_n Haar-distributed orthogonal or unitary matrices with respective sizes n and m_n. In this case, one can prove (with technical additional hypotheses), that almost surely,

$$\lim_{n\to\infty} \tfrac{1}{n} \log \textstyle\int e^{\sqrt{nm_n}\theta\, \Re(\mathrm{Tr}(E_n U_n M_n V_n))} \mathrm{d}U_n \mathrm{d}V_n =$$
$$\lim_{n\to\infty} \tfrac{1}{n} \log \textstyle\int e^{\sqrt{nm_n}\theta\, \Re(\mathrm{Tr}(E_n U_n A_n V_n))} \mathrm{d}U_n \mathrm{d}V_n + \tfrac{1}{n} \log \textstyle\int e^{\sqrt{nm_n}\theta\, \Re(\mathrm{Tr}(E_n U_n B_n V_n))} \mathrm{d}U_n \mathrm{d}V_n.$$

To prove it, the ideas are the same as the ones of [12, Sect. 6.1 and 6.2]: first prove that the random variable

$$\frac{1}{n} \log \int e^{\sqrt{nm_n}\theta\, \Re(\mathrm{Tr}(E_n U_n M_n V_n))} \mathrm{d}U_n \mathrm{d}V_n$$

concentrates around its mean (to do it, use [1, Cor. 4.4.30] instead of [12, Lem. 24] in order to be allowed to consider the complex case) and then the ideas of [12, Sect. 6.2]. Since the singular law of M_n converges almost surely to $\mu_A \boxplus_\lambda \mu_B$, this gives a new proof of (3).

1.3.3. *Strong continuity property for the rectangular spherical integrals.* One other way to extend this work, suggested to us by a referee, would be to use (28) to prove a strong continuity property for the rectangular spherical integrals. in the spirit of Proposition 2.1 or Lemma 2.3 of [17].

2. Preliminaries about the rectangular R-transform

Let μ be a probability measure on the real line which support is contained in $[-K, K]$, with $K > 0$ (we do not suppose μ to be symmetric, how it was the case in the initial definition of the rectangular R-transform). Let us define the generating function of the moments of μ^2

$$M_{\mu^2}(z) = \int_{t\in\mathbb{R}} \frac{t^2 z}{1-t^2 z}\mathrm{d}\mu(t) = \int_{t\in\mathbb{R}} \frac{1}{1-t^2 z}\mathrm{d}\mu(t) - 1 \qquad (z \in [0, K^{-2})).$$

It can easily be proved that M_{μ^2} is nonnegative and non decreasing on $[0, K^{-2})$. Let us define, for $\lambda \in [0,1]$, $T^{(\lambda)}(z) = (\lambda z + 1)(z+1)$, and

$$H_\mu^{(\lambda)}(z) = zT^{(\lambda)}(M_{\mu^2}(z)).$$

Then $H_\mu^{(\lambda)}$ defines an increasing analytic diffeomorphism (in this paper, for I an interval, we shall call an *analytic function on* I a function on I which extends analytically to an open subset of $\mathbb{C}$ containing I) from $[0, K^{-2})$ onto the (possibly unbounded) interval $[0, \lim_{z\uparrow K^{-2}} H_\mu^{(\lambda)}(z))$ such that

$$H_\mu^{(\lambda)}(0) = 0, \qquad \partial_z H_\mu^{(\lambda)}(0) = 1, \qquad H_\mu^{(\lambda)}(z) \geq z, \qquad \lim_{z\uparrow K^{-2}} H_\mu^{(\lambda)}(z) \geq K^{-2}.$$

We denote its inverse by ${H_\mu^{(\lambda)}}^{-1}$. Moreover, $T^{(\lambda)}$ defines an analytic increasing diffeomorphism from $[-1, +\infty)$ to $[0, +\infty)$, thus one can define the *rectangular R-transform with ratio* λ of μ:

$$C_\mu^{(\lambda)}(z) = {T^{(\lambda)}}^{-1}\left(\frac{z}{{H_\mu^{(\lambda)}}^{-1}(z)}\right) \text{ for } z \neq 0, \quad \text{and} \quad C_\mu^{(\lambda)}(0) = 0, \tag{9}$$

which is analytic and non negative on the interval $[0, \lim_{z\uparrow K^{-2}} H_\mu^{(\lambda)}(z))$ (which always contains $[0, K^{-2})$).

By Theorems 3.8 and 3.12 of [4], the rectangular R-transform characterizes symmetric measures, and for all pair μ_1, μ_2 of compactly supported symmetric probability measures, $\mu_1 \boxplus_\lambda \mu_2$ is characterized by the fact that in a neighborhood of zero,

$$C^{(\lambda)}_{\mu_1\boxplus_\lambda\mu_2}(z) = C^{(\lambda)}_{\mu_1}(z) + C^{(\lambda)}_{\mu_2}(z).$$

The following theorem states the continuity of the mapping $(\lambda, \mu) \longmapsto C_\mu^{(\lambda)}$ in a way which is quite different from the one of Theorem 3.11 of [4] (where λ was fixed).

Theorem 2.1. *Fix $K > 0$, let μ_n be a sequence of probability measures on $[-K, K]$ which converges weakly to a limit μ, and let λ_n be a sequence of elements of $[0,1]$ which converges* *to a limit $\lambda \in [0,1]$. Then the sequence of functions $C_{\mu_n}^{(\lambda_n)}$ converges to $C_{\mu}^{(\lambda)}$ uniformly on every compact subset of $[0, K^{-2})$.*

Proof. Recall that $C_{\mu}^{(\lambda)}$ is defined by (9). Since, by Heine's Theorem, $(\lambda, z) \mapsto {T^{(\lambda)}}^{-1}(z)$ is uniformly continuous on every compact subset of $[0,1] \times [0, +\infty)$, it suffices to prove that $\frac{z}{{H_{\mu_n}^{(\lambda_n)}}^{-1}(z)}$ converges to $\frac{z}{{H_{\mu}^{(\lambda)}}^{-1}(z)}$ uniformly on every compact subset of $[0, K^{-2})$.
Claim a : *For each compact subset E of $\mathbb{C}\backslash[K^{-2}, +\infty)$, there is a constant k_E such that for any law ν on $[-K, K]$, for any $c \in [0,1]$, for any $z \in E$,*

$$|T^{(c)}(M_{\nu^2}(z))| \leq k_E.$$

Indeed, for $z \in \mathbb{C}\backslash[K^{-2}, +\infty)$, for any law ν on $[-K, K]$, for any $c \in [0,1]$,

$$T^{(c)}(M_{\nu^2}(z)) = \int_{(t,t')\in[-K,K]^2} \frac{1}{(1-zt^2)(1-zt'^2)} \mathrm{d}\nu(t)\mathrm{d}(c\nu + (1-c)\delta_0)(t'),$$

thus $k_E = \max\{|1 - zt^2|^{-2}\,;\, |t| \leq K, z \in E\}$ is convenient.
Claim b : *As ν varies in the set of laws on $[-K, K]$ and c varies in $[0,1]$, the set of functions*

$$z \in [0, K^{-2}) \longmapsto \frac{z}{{H_{\nu}^{(c)}}^{-1}(z)}$$

is relatively compact for the topology of uniform convergence on every compact subset of $[0, K^{-2})$. By Ascoli's Theorem, to prove Claim b, it suffices to prove that this family is uniformly bounded and uniformly Lipschitz on every compact subset of $[0, K^{-2})$. Let us fix ν a law on $[-K, K]$ and $c \in [0,1]$. Note that we have

$$\frac{z}{{H_{\nu}^{(c)}}^{-1}(z)} = \frac{H_{\nu}^{(c)}(z)}{z} \circ {H_{\nu}^{(c)}}^{-1}(z), \quad \partial_z \frac{z}{{H_{\nu}^{(c)}}^{-1}(z)} = \frac{z{H_{\nu}^{(c)}}'(z) - H_{\nu}^{(c)}(z)}{z^2 {H_{\nu}^{(c)}}'(z)} \circ {H_{\nu}^{(c)}}^{-1}(z).$$

Since, moreover, for all $z \in [0, K^{-2})$, ${H_{\nu}^{(c)}}^{-1}(z) \leq z$ (indeed, for all $z \in [0, K^{-2})$, $H_{\nu}^{(c)}(z) \geq z$), it suffices to verify that the sets of functions

$$\{z \longmapsto \tfrac{H_{\nu}^{(c)}(z)}{z}\,;\, \nu \text{ law on } [-K,K],\, c \in [0,1]\}$$
$$\text{and } \{z \longmapsto \tfrac{z{H_{\nu}^{(c)}}'(z) - H_{\nu}^{(c)}(z)}{z^2 {H_{\nu}^{(c)}}'(z)}\,;\, \nu \text{ law on } [-K,K],\, c \in [0,1]\}$$

are uniformly bounded on every compact subset of $[0, K^{-2})$. The family of functions $\frac{H_{\nu}^{(c)}(z)}{z} = T^{(c)}(M_{\nu^2}(z))$, indexed by ν, c, is a family of analytic functions on $\mathbb{C}\backslash[K^{-2}, +\infty)$ which is uniformly bounded on every compact subset of $\mathbb{C}\backslash[K^{-2}, +\infty)$ (by Claim a). As a consequence, the family of the derivatives $\partial_z \frac{H_{\nu}^{(c)}(z)}{z}$ is also uniformly bounded on every compact subset of $\mathbb{C}\backslash[K^{-2}, +\infty)$. Since

$$\frac{z{H_{\nu}^{(c)}}'(z) - H_{\nu}^{(c)}(z)}{z^2 {H_{\nu}^{(c)}}'(z)} = \frac{1}{{H_{\nu}^{(c)}}'(z)} \partial_z \frac{H_{\nu}^{(c)}(z)}{z}$$

and ${H_{\nu}^{(c)}}'(z) \geq 1$ on $[0, K^{-2})$, Claim b is proved.

Hence one can suppose that $\frac{z}{H_{\mu_n}^{(\lambda_n)^{-1}}(z)}$ converges to a function f uniformly on every compact of $[0, K^{-2})$. Let us fix $z \in [0, K^{-2})$ and let us prove that $f(z) = \frac{z}{H_{\mu}^{(\lambda)^{-1}}(z)}$. If $z = 0$, it is clear (since all these functions are implicitly defined to map 0 to 1). Suppose that $z > 0$. Note that $f(z) \neq 0$, because for all n, $\frac{z}{H_{\mu_n}^{(\lambda_n)^{-1}}(z)} \geq 1$. Let us denote $l = \frac{z}{f(z)}$. It suffices to prove that $l = H_{\mu}^{(\lambda)^{-1}}(z)$, i.e. that $H_{\mu}^{(\lambda)}(l) = z$. Since

$$H_{\mu}^{(\lambda)}(l) = \lim_{n\to\infty} H_{\mu}^{(\lambda)}(H_{\mu_n}^{(\lambda_n)^{-1}}(z)),$$

it suffices to prove that $H_{\mu_n}^{(\lambda_n)}$ converges to $H_{\mu}^{(\lambda)}$ uniformly on every compact subset of $[0, K^{-2})$. But it is easy to see, using [1, Th. C.11], that $M_{\mu_n^2}$ converges to M_{μ^2} uniformly on every compact subset of $[0, K^{-2})$ and then that $H_{\mu}^{(\lambda_n)}$ converges to $H_{\mu}^{(\lambda)}$ uniformly on every compact subset of $[0, K^{-2})$. □

3. Proof of Theorem 1.2

3.1. **Preliminaries.** We shall use the following lemmas several times in the paper. Let $\|\cdot\|$ denote the canonical euclidian norm on each $\mathbb{R}^d$.

Lemma 3.1. *Let $(G_i)_{i\geq 1}$ be a family of independent real random variables with standard Gaussian law. Let T be fixed and let, for each n, $(\sigma_{n,1}, \ldots, \sigma_{n,n}) \in [0, T]^n$ be such that*

$$\frac{1}{n}\sum_{i=1}^{n} \sigma_{n,i}^2 = 1. \tag{10}$$

Let us define, for each n, $X_n = (\sigma_{n,1}G_1, \ldots, \sigma_{n,n}G_n)$. Then for all $\kappa \in (0, \frac{1}{2})$, for n large enough,

$$\mathbb{P}\{|\|X_n\| - \sqrt{n}| \geq n^{\frac{1}{2}-\kappa}\} \leq 2T^4 n^{2\kappa-1}.$$

Proof. Note that by (10), the random variable $N_n := \frac{\|X_n\|^2}{n} - 1$ is centered. Moreover, $\operatorname{Var}(N_n) = \frac{\operatorname{Var}(G_1^2)}{n^2}\sum_{i=1}^n \sigma_{n,i}^4 \leq \frac{2T^4}{n}$. It follows, by Tchebichev's inequality, that for all $\kappa \in (0, \frac{1}{2})$,

$$\mathbb{P}\{|N_n| \geq n^{-\kappa}\} \leq 2T^4 n^{2\kappa-1}.$$

To deduce that for n large enough,

$$\mathbb{P}\left\{\left|\frac{\|X_n\|}{\sqrt{n}} - 1\right| \geq n^{-\kappa}\right\} \leq 2T^4 n^{2\kappa-1},$$

it suffices to notice that the function $\sqrt{\cdot}$ is 1-Lipschitz on $[1/4, +\infty)$ and that $n^{-\kappa} \leq 3/4$ for n large enough. □

Lemma 3.2. *Let μ be a probability measure whose support is contained in $[-K, K]$, fix $\lambda \in [0, 1]$, $\theta \in [0, K^{-1})$ and define $\gamma = C_{\mu}^{(\lambda)}(\theta^2)$. Then*

$$M_{\mu^2}\left(\frac{\theta^2}{T^{(\lambda)}(\gamma)}\right) = \gamma. \tag{11}$$

Proof. By the definition of $C_\mu^{(\lambda)}$ given in (9), $\frac{\theta^2}{{H_\mu^{(\lambda)}}^{-1}(\theta^2)} = T^{(\lambda)}(\gamma)$, hence $\frac{\theta^2}{T^{(\lambda)}(\gamma)} = {H_\mu^{(\lambda)}}^{-1}(\theta^2)$. Since $\gamma \geq 0$, $\frac{\theta^2}{T^{(\lambda)}(\gamma)} \in [0, K^{-2})$ and one can apply the function $H_\mu^{(\lambda)}$ on both sides. We get $H_\mu^{(\lambda)}\left(\frac{\theta^2}{T^{(\lambda)}(\gamma)}\right) = \theta^2$, i.e.

$$\frac{\theta^2}{T^{(\lambda)}(\gamma)} T^{(\lambda)}\left(M_{\mu^2}\left(\frac{\theta^2}{T^{(\lambda)}(\gamma)}\right)\right) = \theta^2.$$

It follows that $T^{(\lambda)}\left(M_{\mu^2}\left(\frac{\theta^2}{T^{(\lambda)}(\gamma)}\right)\right) = T^{(\lambda)}(\gamma)$. Since both $M_{\mu^2}\left(\frac{\theta^2}{T^{(\lambda)}(\gamma)}\right)$ and γ are non-negative real numbers, one gets (11). □

Lemma 3.3. *Let X_n be a sequence of nonnegative random variables, with positive expectations. Let Z_n be a sequence of real random variables such that there exists deterministic constants $C, \eta > 0$ such that for all n, $|Z_n| \leq Cn^{1-\eta}$. Then as n tends to infinity,*

$$\frac{1}{n} \log \mathbb{E}(X_n e^{Z_n}) = \frac{1}{n} \log \mathbb{E}(X_n) + o(1).$$

Proof. It suffices to notice that we have $X_n e^{-Cn^{1-\eta}} \leq X_n e^{Z_n} \leq X_n e^{Cn^{1-\eta}}$. □

3.2. **Notation for the proof of Theorem 1.2.** In the next sections, $o(1)$ shall denote any sequence of functions on $(-K^{-1}, K^{-1})$ which converges to zero as n tends to infinity, uniformly on every compact subset of $(-K^{-1}, K^{-1})$. Also, we shall work with the notation introduced in Remark 1.1 and handle $I_n(\theta)$ via Formula (7).

3.3. **Reduction to the case where all singular values of M_n are positive.** Let us suppose the result to be proved in the particular case where for all n, k, $\mu_{n,k} > 0$, and let us prove it in the general case. We set, for each n, k,

$$\tilde{\mu}_{n,k} = \mu_{n,k} + \min\{m_n^{-2}, (K - \mu_{n,k})/2\}$$

and define the perturbation of $I_n(\theta)$

$$\tilde{I}_n(\theta) = \frac{1}{n} \log \mathbb{E}\{\exp[\sqrt{nm_n}\theta \sum_{k=1}^{n} \Re(u_{n,k}\tilde{\mu}_{n,k}v_{n,k})]\}.$$

The uniform law on the $\tilde{\mu}_{n,k}$'s converges weakly to μ as n tends to infinity and we have $0 < \tilde{\mu}_{n,k} < K$, so by hypothesis, if follows that $\lim_{n\to\infty} \tilde{I}_n(\theta) = I(\theta)$. Note moreover that

$$\tilde{I}_n(\theta) = \frac{1}{n} \log \mathbb{E}\{\exp[\sqrt{nm_n}\theta \sum_{k=1}^{n} \Re(u_{n,k}\mu_{n,k}v_{n,k})]e^{Z_n}\},$$

with

$$Z_n = \sqrt{nm_n}\theta \sum_{k=1}^{n} \Re(u_{n,k}v_{n,k}) \min\{m_n^{-2}, (K - \mu_{n,k})/2\}.$$

Since $|Z_n| \leq K^{-1}$, Lemma 3.3 allows us to claim that $\lim_{n\to\infty} I_n(\theta) = I(\theta)$.

3.4. **Deducing the complex case from the real one.** Let us explain how one can deduce the complex case from the real one. Let us use, in this paragraph, the notation $I^{(\beta)}_{n,m_n}(\theta, M_n)$ to emphasize on the value of each of the parameters defining $I_n(\theta)$. We have

$$\begin{aligned}
I^{(2)}_{n,m_n}(\theta, M_n) &= \frac{1}{n}\log \mathbb{E}\{\exp[\sqrt{nm_n}\theta\sum_{k=1}^{n}\Re(u_{n,k}\mu_{n,k}v_{n,k})]\}\\
&= \frac{1}{n}\log \mathbb{E}\{\exp[\sqrt{nm_n}\theta\sum_{k=1}^{n}\mu_{n,k}(\Re(u_{n,k})\Re(v_{n,k}) - \Im(u_{n,k})\Im(v_{n,k}))]\}\\
&= 2I^{(1)}_{2n,2m_n}\left(\frac{\theta}{2}, \begin{bmatrix} M_n & 0\\ 0 & -M_n\end{bmatrix}\right) \qquad (12)
\end{aligned}$$

Indeed, for $(z_{n,1}, \ldots, z_{n,n})$ a vector with uniform distribution on the unit sphere of $\mathbb{C}^n$ the vector $(\Re(z_1), \Im(z_1), \ldots, \Re(z_n), \Im(z_n))$ has uniform distribution on the unit sphere of $\mathbb{R}^{2n}$ (use the realization of such vectors as the projection of standard Gaussian vectors on the sphere to see it).

Since the singular values of $\begin{bmatrix} M_n & 0\\ 0 & -M_n\end{bmatrix}$ are the one of M_n with multiplicity multiplied by two, (12) allows to deduce the complex case from the real one.

Let us now prove Theorem 1.2 in the case where all $\mu_{n,k}$'s are positive and where $\beta = 1$.

3.5. **Proof of Theorem 1.2: a) Expression of $I_n(\theta)$ with a Gaussian integral.** As written above, we suppose from now on that all $\mu_{n,k}$'s are positive and that $\beta = 1$. In this section, we shall first explain how to replace, in Formula (7), $\sqrt{n}u_{n,k}$ and $\sqrt{m_n}v_{n,k}$ by independent standard Gaussian variables (Formula (17)) and then inject (quite artificially first) $C^{(\lambda)}_{\mu}(\theta^2)$ in the formula of $I_n(\theta)$ (Equation (21)).

For each n, let us define the function

$$f_n : ((x_1, \ldots, x_n), (y_1, \ldots, y_{m_n})) \in \mathbb{R}^n \times \mathbb{R}^{m_n} \longmapsto \sum_{k=1}^{n} \mu_{n,k}x_k y_k.$$

By [12, Sect. 1.2.1], up to a change of the probability space which does not change the expectation, one can suppose that there are independent standard Gaussian vectors x_n, y_n of $\mathbb{R}^n$, respectively $\mathbb{R}^{m_n}$, such that

$$u_n = \frac{x_n}{\|x_n\|}, \qquad v_n = \frac{y_n}{\|y_n\|}.$$

Let us fix $\kappa \in (0, 1/2)$. If $\lambda > 0$, the precise choice of $\kappa \in (0, 1/2)$ is irrelevant, but if $\lambda = 0$, we choose $\kappa \in (\frac{\alpha-1}{2}, \frac{1}{2})$ (α is the one of (6)). Let us now define the set

$$A_n := \left\{(x, y) \in \mathbb{R}^n \times \mathbb{R}^{m_n}\ ;\ \left|\|x\| - \sqrt{n}\right| \le n^{\frac{1}{2}-\kappa}, \left|\|y\| - \sqrt{m_n}\right| \le m_n^{\frac{1}{2}-\kappa}\right\}.$$

The event $\{(x_n, y_n) \in A_n\}$ is independent of (u_n, v_n) (because the density of the law of a standard Gaussian vector is a radial function), thus

$$I_n(\theta) \quad = \quad \frac{1}{n} \log \mathbb{E}\left[\mathbb{1}_{(x_n,y_n)\in A_n} \exp(\sqrt{nm_n}\theta f_n(u_n, v_n))\right] - \frac{1}{n} \log \mathbb{P}(A_n).$$

Moreover, by Lemma 3.1, $\mathbb{P}\{(x_n, y_n) \in A_n\} \longrightarrow 1$ as $n \to \infty$, thus

$$I_n(\theta) = \frac{1}{n} \log \mathbb{E}\left[\mathbb{1}_{(x_n,y_n)\in A_n} \exp(\sqrt{nm_n}\theta f_n(u_n, v_n))\right] + o(1). \tag{13}$$

Moreover, note that on the event $\{(x_n, y_n) \in A_n\}$,

$$\begin{aligned} \sqrt{n} - n^{\frac{1}{2}-\kappa} \leq \quad & \|x_n\| \quad \leq \sqrt{n} + n^{\frac{1}{2}-\kappa} \\ \sqrt{m_n} - m_n^{\frac{1}{2}-\kappa} \leq \quad & \|y_n\| \quad \leq \sqrt{m_n} + m_n^{\frac{1}{2}-\kappa}, \end{aligned}$$

thus, since $m_n \geq n$,

$$\sqrt{nm_n} - 3\sqrt{m_n}n^{\frac{1}{2}-\kappa} \leq \|x_n\|\|y_n\| \leq \sqrt{nm_n} + 3\sqrt{m_n}n^{\frac{1}{2}-\kappa}. \tag{14}$$

If $\lambda > 0$, since m_n/n is bounded, it follows that there is a deterministic constant C independent of n such that on on the event $\{(x_n, y_n) \in A_n\}$,

$$|\, \|x_n\|\|y_n\| - \sqrt{nm_n}\,| \leq Cn^{1-\kappa}. \tag{15}$$

If $\lambda = 0$, it follows from (14) and (6) that for $\eta = \frac{1-\alpha}{2} + \kappa$ (which is positive by definition of κ), for n large enough,

$$|\, \|x_n\|\|y_n\| - \sqrt{nm_n}\,| \leq 3n^{1-\eta}. \tag{16}$$

Note that by (13),

$$I_n(\theta) = \frac{1}{n} \log \mathbb{E}[\mathbb{1}_{(x_n,y_n)\in A_n} e^{\theta f_n(x_n,y_n) + \frac{\theta f_n(x_n,y_n)}{\|x_n\|\|y_n\|}(\sqrt{nm_n} - \|x_n\|\|y_n\|)}] + o(1),$$

and that for all n, k, $0 \leq \mu_{k,n} \leq K$, which implies that $\left|\frac{f_n(x_n,y_n)}{\|x_n\|\|y_n\|}\right| \leq K$. Hence by Lemma 3.3 and (15) (or (16) if $\lambda = 0$),

$$I_n(\theta) = \frac{1}{n} \log \mathbb{E}\left[\mathbb{1}_{(x_n,y_n)\in A_n} e^{\theta f_n(x_n,y_n)}\right] + o(1). \tag{17}$$

Note that on the event $\{(x_n, y_n) \in A_n\}$, we have

$$\begin{aligned} n - 2n^{1-\kappa} \leq n - 2n^{1-\kappa} + n^{1-2\kappa} \leq \quad & \|x_n\|^2 \quad \leq n + 2n^{1-\kappa} + n^{1-2\kappa} \leq n + 3n^{1-\kappa} \\ n - 2n^{1-\kappa} \leq n - 2nm_n^{-\kappa} + nm_n^{-2\kappa} \leq \quad & \tfrac{n}{m_n}\|y_n\|^2 \quad \leq n + 2nm_n^{-\kappa} + nm_n^{-2\kappa} \leq n + 3n^{1-\kappa}. \end{aligned}$$

thus for all n, on the event $\{(x_n, y_n) \in A_n\}$,

$$\left|\, \|x_n\|^2 - n\,\right| + \left|\, \frac{n}{m_n}\|y_n\|^2 - n\,\right| \leq 6n^{1-\kappa}. \tag{18}$$

Now, let us define, for each n,

$$\gamma_n(\theta) = C_{\mu_n}^{(\frac{n}{m_n})}(\theta^2) \qquad \text{for } \mu_n = \frac{1}{n}\sum_{k=1}^{n} \delta_{\mu_{n,k}}. \tag{19}$$

Note that μ_n is the singular law of M_n, which tends to μ. Hence by Theorem 2.1, we have

$$\gamma_n(\theta) \underset{n\to\infty}{\longrightarrow} C_\mu^{(\lambda)}(\theta^2) \qquad \text{uniformly on every compact subset of } (-K^{-1}, K^{-1}), \tag{20}$$

so by (18), for every such compact set E, there is a constant Q_E such that for all n, for all $\theta \in E$, on the event $\{(x_n, y_n) \in A_n\}$, we have

$$|\gamma_n(\theta)(\frac{1}{2}\|x_n\|^2 + \frac{n}{2m_n}\|y_n\|^2 - n)| \le Q_E n^{1-\kappa}.$$

Hence, by (17) and Lemma 3.3,

$$\begin{aligned} I_n(\theta) =& \frac{1}{n}\log \mathbb{E}\left[\mathbb{1}_{(x_n,y_n)\in A_n} \exp\left\{\theta f_n(x_n,y_n) - \gamma_n(\theta)\left(\frac{1}{2}\|x_n\|^2 + \frac{n}{2m_n}\|y_n\|^2 - n\right)\right\}\right] + o(1) \\ =& \gamma_n(\theta) + \frac{1}{n}\log \underbrace{\mathbb{E}\left[\mathbb{1}_{(x_n,y_n)\in A_n} \exp\left\{\theta f_n(x_n,y_n) - \gamma_n(\theta)\left(\frac{1}{2}\|x_n\|^2 + \frac{n}{2m_n}\|y_n\|^2\right)\right\}\right]}_{:=J_n(\theta)} + o(1). \end{aligned}$$

Thus, by (20),

$$I_n(\theta) = C_\mu^{(\lambda)}(\theta^2) + \frac{1}{n}\log J_n(\theta) + o(1). \tag{21}$$

3.6. **Proof of Theorem 1.2: b) Convergence of the Gaussian integral.** We have, assimilating the vectors of $\mathbb{R}^n$ and $\mathbb{R}^{m_n}$ with column-matrices,

$$J_n(\theta) = (2\pi)^{-\frac{n+m_n}{2}} \int_{x\in\mathbb{R}^n, y\in\mathbb{R}^{m_n}} \mathbb{1}_{A_n}(x,y) \exp\left\{-\frac{1}{2}\begin{bmatrix} x^t & y^t\end{bmatrix} T_n \begin{bmatrix} x \\ y \end{bmatrix}\right\} \mathrm{d}x\mathrm{d}y, \tag{22}$$

for

$$T_n := \begin{bmatrix} a_n(\theta)I_n & \Lambda_n(\theta) & 0_{n,m_n-n} \\ \Lambda_n(\theta) & b_n(\theta)I_n & 0_{n,m_n-n} \\ 0_{m_n-n,n} & 0_{m_n-n,n} & b_n(\theta)I_{m_n-n} \end{bmatrix},$$

where $a_n(\theta) = 1+\gamma_n(\theta)$, $b_n(\theta) = 1 + \frac{n}{m_n}\gamma_n(\theta)$ and $\Lambda_n(\theta)$ is the diagonal $n\times n$ matrix with diagonal entries

$$\lambda_{n,1}(\theta) := -\theta\mu_{n,1}, \ldots, \lambda_{n,n}(\theta) := -\theta\mu_{n,n}.$$

Notation: In this section, in order to lighten the notation, we shall write J_n for $J_n(\theta)$, a_n for $a_n(\theta)$, *etc.* We shall also use the notation of the matricial functional calculus (thus assimilate a and aI_n, etc.).

Lemma 3.4. *Let us fix $n \ge 1$ and let a, b be real numbers and Λ an invertible diagonal real $n\times n$ matrix. Let us define*

$$\Delta = (b-a)^2 + 4\Lambda^2, \quad r^\pm = \frac{a+b\pm\sqrt{\Delta}}{2}, \quad f^\pm = \frac{1}{\sqrt{2\Delta \pm 2(b-a)\sqrt{\Delta}}}$$

and

$$T = \begin{bmatrix} a & \Lambda \\ \Lambda & b\end{bmatrix}, \quad D = \begin{bmatrix} r^+ & 0 \\ 0 & r^-\end{bmatrix}, \quad P = \begin{bmatrix} 2\Lambda f^+ & 2\Lambda f^- \\ (b-a)f^+ + \sqrt{\Delta}f^+ & (b-a)f^- - \sqrt{\Delta}f^-\end{bmatrix}.$$

Then P is an orthogonal matrix and we have $T = PDP^t$.

Proof. One can easily verify that P is orthogonal. Let us define

$$Q = \begin{bmatrix} 2\Lambda & 2\Lambda \\ (b-a)+\sqrt{\Delta} & (b-a)-\sqrt{\Delta} \end{bmatrix}, H = \begin{bmatrix} f^+ & 0 \\ 0 & f^- \end{bmatrix}.$$

Then $P = QH$. One can easily verify that $TQ = QD$. If follows that $TQH = QDH$. Since $HD = DH$, $TQH = QHD$, i.e. $TP = PD$, thus $T = PDP^t$. □

For $\theta \neq 0$, let us define Δ_n, $r_n^\pm$, $f_n^\pm$ as in the lemma, using Λ_n instead of Λ, a_n instead of a and b_n instead of b. Let us define P_n in the same way, extended to an $(n+m_n)\times(n+m_n)$ matrix by adding I_{m_n-n} on the lower-right corner, i.e.

$$P_n = \begin{bmatrix} 2\Lambda_n f_n^+ & 2\Lambda_n f_n^- & 0 \\ (b_n-a_n)f_n^+ + \sqrt{\Delta_n} f_n^+ & (b_n-a_n)f_n^- - \sqrt{\Delta_n} f^- & 0 \\ 0 & 0 & I_{m_n-n} \end{bmatrix},$$

and D_n extended to an $(n+m_n)\times(n+m_n)$ matrix by adding $b_n I_{m_n-n}$ on the lower-right corner, i.e.

$$D_n = \begin{bmatrix} r_n^+ & 0 & 0 \\ 0 & r_n^- & 0 \\ 0 & 0 & b_n I_{m_n-n} \end{bmatrix}.$$

For $\theta = 0$, we set $r_n^\pm = 1$, $P_n = D_n = I_{n+m_n}$.

Let us denote, for X an $(n+m_n)\times(n+m_n)$ matrix, $X(A_n) = \{X \begin{bmatrix} x \\ y \end{bmatrix} \,;\, (x,y) \in A_n\}$. Let us also introduce a standard Gaussian random column vector in $\mathbb{R}^{n+m_n}$, that we shall denote by

$$Z_n = (\underbrace{Z_{n,1}^+,\ldots,Z_{n,n}^+}_{:=Z_n^+}, \underbrace{Z_{n,1}^-,\ldots,Z_{n,n}^-}_{:=Z_n^-}, \underbrace{Z_{n,1}^0,\ldots,Z_{n,m_n-n}^0}_{:=Z_n^0})^t.$$

We have, by (22) and Lemma 3.4,

$$J_n = (2\pi)^{-\frac{n+m_n}{2}} \int_{A_n} \exp\{-\frac{1}{2}\begin{bmatrix} x^t & y^t \end{bmatrix} P_n D_n P_n^t \begin{bmatrix} x \\ y \end{bmatrix}\} \mathrm{d}x \mathrm{d}y.$$

Thus, since P_n is an orthogonal matrix,

$$J_n = (2\pi)^{-\frac{n+m_n}{2}} \int_{P_n^t(A_n)} \exp\{-\frac{1}{2}\begin{bmatrix} x^t & y^t \end{bmatrix} D_n \begin{bmatrix} x \\ y \end{bmatrix}\} \mathrm{d}x \mathrm{d}y.$$

Hence, by definition of D_n, we have

$$J_n = (2\pi)^{-\frac{n+m_n}{2}} [b_n^{m_n-n} \prod_{i=1}^n r_{n,i}^+ r_{n,i}^-]^{-1/2} \int_{\sqrt{D_n} P_n^t(A_n)} \exp\{-\frac{1}{2}(\|x\|^2 + \|y\|^2)\} \mathrm{d}x \mathrm{d}y,$$

which, by definition of Z_n, can be written

$$J_n = [b_n^{m_n-n} \prod_{i=1}^n (a_n b_n - \lambda_{n,i}^2)]^{-1/2} \underbrace{\mathbb{P}\{Z_n \in \sqrt{D_n} P_n^t(A_n)\}}_{=\mathbb{P}\{P_n D_n^{-1/2} Z_n \in A_n\}} \tag{23}$$

Let $X_n = (X_{n,1}, \ldots, X_{n,n})^t$ be the vector of the first n coordinates of $P_n D_n^{-1/2} Z_n$ and $Y_n = (Y_{n,1}, \ldots, Y_{n,m_n})^t$ be the one of the m_n last ones. By definition of the set A_n, we have

$$P_n D_n^{-1/2} Z_n \in A_n \iff \left|\|X_n\| - \sqrt{n}\right| \le n^{-\kappa} \text{ and } \left|\|Y_n\| - \sqrt{m_n}\right| \le m_n^{-\kappa}. \tag{24}$$

<u>Claim</u> : *Both events* $\{|\|X_n\| - \sqrt{n}| \le n^{-\kappa}\}$ *and* $\left\{\left|\|Y_n\| - \sqrt{m_n}\right| \le m_n^{-\kappa}\right\}$ *have probabilities tending to one as* n *tends to infinity, uniformly on every compact subset of* $(-K^{-1}, K^{-1})$ *(the random vectors* X_n *and* Y_n *depend indeed on* θ*).*

For $\theta = 0$, X_n (resp. Y_n) is a standard Gaussian random vector of $\mathbb{R}^n$ (resp. $\mathbb{R}^{m_n}$). For $\theta \neq 0$, by the definitions of P_n and D_n,

$$\begin{aligned} X_n &= 2\Lambda_n (f_n^+ (r_n^+)^{-1/2} Z_n^+ + f_n^- (r_n^-)^{-1/2} Z_n^-), \\ Y_n &= \begin{bmatrix} (b_n - a_n + \sqrt{\Delta_n}) f_n^+ (r_n^+)^{-1/2} Z_n^+ + (b_n - a_n - \sqrt{\Delta_n}) f_n^- (r_n^-)^{-1/2} Z_n^- \\ b_n^{-1/2} Z_n^0 \end{bmatrix}. \end{aligned}$$

Thus for each n, X_n (resp. Y_n) has the law of

$$(\sigma_{n,1} G_1, \ldots, \sigma_{n,n} G_n) \qquad (\text{resp. } (\sigma'_{n,1} G_1, \ldots, \sigma'_{n,m_n} G_{m_n})),$$

for $(G_i)_{i \ge 1}$ a family of independent real random variables with standard Gaussian law and where if $\theta = 0$, all $\sigma_{n,i}$'s and $\sigma'_{n,i}$'s are equal to 1, whereas if $\theta \neq 0$, for each $i = 1, \ldots, n$,

$$\begin{aligned} \sigma_{n,i}^2 &= 4\lambda_{n,i}^2 \frac{2}{(2\Delta_{n,i} + 2(b_n - a_n)\sqrt{\Delta_{n,i}})(a_n + b_n + \sqrt{\Delta_{n,i}})} \\ &\qquad + 4\lambda_{n,i}^2 \frac{2}{(2\Delta_{n,i} - 2(b_n - a_n)\sqrt{\Delta_{n,i}})(a_n + b_n - \sqrt{\Delta_{n,i}})}, \end{aligned}$$

for each $i = n+1, \ldots, m_n$, ${\sigma'_{n,i}}^2 = b_n^{-1}$ and for each $i = 1, \ldots, n$,

$${\sigma'_{n,i}}^2 = \frac{2(b_n - a_n + \sqrt{\Delta_n})^2}{(a_n + b_n + \sqrt{\Delta_n})[2(b_n - a_n)\sqrt{\Delta_n} + 2\Delta_n]} + \frac{2(\sqrt{\Delta_n} - (b_n - a_n))^2}{(a_n + b_n - \sqrt{\Delta_n})[-2(b_n - a_n)\sqrt{\Delta_n} + 2\Delta_n]}.$$

Hence by Lemma 3.1, to prove the claim, it suffices to prove:

$$\forall \varepsilon > 0, \quad \sup_{|\theta| \le K^{-1} - \varepsilon} \sup_{\substack{1 \le i \le n \\ n \ge 1}} \sigma_{n,i} < +\infty \quad \text{and} \quad \frac{1}{n} \sum_{i=1}^{n} \sigma_{n,i}^2 = 1, \tag{25}$$

$$\forall \varepsilon > 0, \quad \sup_{|\theta| \le K^{-1} - \varepsilon} \sup_{\substack{1 \le i \le m_n \\ n \ge 1}} \sigma'_{n,i} < +\infty \quad \text{and} \quad \frac{1}{m_n} \sum_{i=1}^{m_n} {\sigma'_{n,i}}^2 = 1. \tag{26}$$

Note first that (25) and (26) both hold when $\theta = 0$.

A straightforward computation leads, for $\theta \neq 0$, to the formula

$$\sigma_{n,i}^2 = \frac{-16 b_n \lambda_{n,i}^2}{(b_n^2 + \Delta_{n,i} - a_n^2)^2 - 4 b_n^2 \Delta_{n,i}}.$$

But (removing the indices)

$$\begin{aligned}(b^2+\Delta-a^2)^2-4b^2\Delta &= (2b^2-2ab+4\lambda^2)^2-4b^2(b^2-2ab+a^2+4\lambda^2)\\ &= 16\lambda^4-16ab\lambda^2.\end{aligned}$$

It follows, writing γ_n for $\gamma_n(\theta)$, that

$$\sigma_{n,i}^2=\frac{b_n}{a_nb_n-\lambda_{n,i}^2}=\frac{1}{1+\gamma_n}\times\frac{T^{(\frac{n}{m_n})}(\gamma_n)}{T^{(\frac{n}{m_n})}(\gamma_n)-\theta^2\mu_{n,i}^2}=\frac{1}{1+\gamma_n}\times\frac{1}{1-\frac{\theta^2}{T^{(\frac{n}{m_n})}(\gamma_n)}\mu_{n,i}^2}. \tag{27}$$

By definition of $\gamma_n(\theta)$, we have $\gamma_n(\theta)\geq 0$, hence $T^{(\frac{n}{m_n})}(\gamma_n(\theta))\geq 1$. Since for all n,i, $|\mu_{n,i}|\leq K$, it follows that the first part of (25) holds. Moreover, by the definition of μ_n given in (19), we have

$$\frac{1}{n}\sum_{i=1}^n\sigma_{n,i}^2=\frac{1}{1+\gamma_n(\theta)}M_{\mu_n^2}\left(\frac{\theta^2}{T^{(\frac{n}{m_n})}(\gamma_n(\theta))}\right)+\frac{1}{1+\gamma_n(\theta)}.$$

By (11), it follows that the second part of (25) also holds.

Let us now prove (26). A straightforward computation leads, for $\theta\neq 0$ and $i\leq n$, to the formula

$${\sigma'_{n,i}}^2 = \frac{a_n}{a_nb_n-\lambda_{n,i}^2}.$$

Hence for $\theta\neq 0$ and $i\leq n$,

$${\sigma'_{n,i}}^2 = \frac{1}{1+\frac{n}{m_n}\gamma_n(\theta)}\times\frac{1}{1-\frac{\theta^2}{T^{(\frac{n}{m_n})}(\gamma_n(\theta))}\mu_{n,i}^2},$$

whereas for $i=n+1,\dots,m_n$, ${\sigma'_{n,i}}^2=\frac{1}{1+\frac{n}{m_n}\gamma_n(\theta)}$. The first part of (26) holds for the same reasons as the first part of (25) above. Moreover, writing γ_n for $\gamma_n(\theta)$, we have

$$\frac{1}{m_n}\sum_{i=1}^{m_n}{\sigma'_{n,i}}^2=\frac{n}{m_n(1+\frac{n}{m_n}\gamma_n)}M_{\mu_n^2}\left(\frac{\theta^2}{T^{(\frac{n}{m_n})}(\gamma_n)}\right)+\frac{n}{m_n(1+\frac{n}{m_n}\gamma_n)}+\frac{m_n-n}{m_n(1+\frac{n}{m_n}\gamma_n)}.$$

By (11), it follows that the second part of (26) also holds.

The proof of the claim is complete. As a consequence, by (24), the probability of the event $\{P_nD_n^{-1/2}Z_n\in A_n\}$ tends to one as n tends to infinity, uniformly on every compact subset of $(-K^{-1},K^{-1})$ (remember indeed that the matrices P_n and D_n depend on θ). So by (23), we have, still writing γ_n for $\gamma_n(\theta)$,

$$\frac{1}{n}\log(J_n(\theta))= \tag{28}$$

$$\begin{aligned}&\frac{n-m_n}{2n}\log(1+\frac{n}{m_n}\gamma_n)-\frac{\log(T^{(\frac{n}{m_n})}(\gamma_n))}{2}-\frac{1}{2}\int_{t\in[-K,K]}\log\{1-\frac{\theta^2}{T^{(\frac{n}{m_n})}(\gamma_n)}t^2\}\mathrm{d}\mu_n(t)+o(1)\\ &=-\frac{m_n}{2n}\log(1+\frac{n}{m_n}\gamma_n)-\frac{\log(1+\gamma_n)}{2}-\frac{1}{2}\int_{t\in[-K,K]}\log\{1-\frac{\theta^2}{T^{(\frac{n}{m_n})}(\gamma_n)}t^2\}\mathrm{d}\mu_n(t)+o(1).\end{aligned}$$

By hypothesis, μ_n, which is defined in (19), converges weakly to μ. Using (20) and [1, Th. C.11], one easily sees that, writing γ for $C_\mu^{(\lambda)}(\theta^2)$, we have

$$\frac{1}{n}\log(J_n(\theta)) = -\frac{1}{2\lambda}\log(1+\lambda\gamma) - \frac{\log(1+\gamma)}{2} - \frac{1}{2}\int_{t\in[-K,K]} \log\{1-\frac{\theta^2}{T^{(\lambda)}(\gamma)}t^2\}\mathrm{d}\mu(t) + o(1),$$

where in the case where $\lambda = 0$, $\frac{1}{2\lambda}\log(1+\lambda\gamma)$ has to be understood as $\frac{\gamma}{2}$. By (21), one gets, still writing γ for $C_\mu^{(\lambda)}(\theta^2)$,

(29)

$$I_n(\theta) = \underbrace{\gamma - \frac{1}{2\lambda}\log(1+\lambda\gamma) - \frac{\log(1+\gamma)}{2} - \frac{1}{2}\int_{t\in[-K,K]} \log\{1-\frac{\theta^2}{T^{(\lambda)}(\gamma)}t^2\}\mathrm{d}\mu(t)}_{:=f(\theta)} + o(1).$$

$I(0) = f(0) = 0$ (indeed, by (9), $C_\mu^{(\lambda)}(0) = 0$). So to conclude the proof of Theorem 1.2, it suffices to verify that I and f have the same derivatives on $(-K^{-1}, K^{-1})$. Using (11), it can easily be proved that both derivatives are equal to γ/θ.

Aknowledgements: The author would like to thank Mylène Maïda for her comments on a draft of this paper and also both referees for their careful reading of the paper and their remarks.

References

[1] G. Anderson, A. Guionnet, O. Zeitouni *An Introduction to Random Matrices.* Cambridge studies in advanced mathematics, 118 (2009).

[2] F. Benaych-Georges *Rectangular random matrices, related free entropy and free Fisher's information.* J. Operator Theory, Vol. 62, no. 2 (2009) 371–419.

[3] F. Benaych-Georges *Infinitely divisible distributions for rectangular free convolution: classification and matricial interpretation.* Probab. Theory Related Fields, Vol. 139, issue 1–2 (2007), 143–189.

[4] F. Benaych-Georges *Rectangular random matrices, related convolution* Probab. Theory Related Fields, Vol. 144, issue 3 (2009), 471–515.

[5] F. Benaych-Georges *On a surprising relation between the Marchenko-Pastur law, rectangular and square free convolutions*, Ann. Inst. Henri Poincaré Probab. Stat., Vol. 46, no. 3 (2010), 644–652.

[6] B. Collins *Moments and cumulants of polynomial random variables on unitary groups, the Itzykson-Zuber integral and free probability.* Int. Math. Res. Not. 17 (2003), 953–982.

[7] B. Collins, P. Śniady *Integration with respect to the Haar measure on unitary, orthogonal and symplectic group* Comm. Math. Phys. 264 (2006), no. 3, 773–795.

[8] B. Collins, P. Śniady *New scaling of Itzykson-Zuber integrals* Annales de l'IHP Prob. Stats. 43, 2 (2007), 139–146.

[9] B. Collins, J.A. Mingo, P. Śniady, R. Speicher *Second order freeness and fluctuations of random matrices. III. Higher order freeness and free cumulants.* Doc. Math. 12 (2007), 1–70.

[10] A. Ghaderipoor, C. Tellambura *Generalization of some integrals over unitary matrices by character expansion of groups* J. Math. Physics, 49 073519 (2008).

[11] A. Guionnet, O. Zeitouni *Large deviations asymptotics for spherical integrals* J. Funct. Anal. 188, 2 (2002), 461–515.

[12] A. Guionnet, M. Maïda *A Fourier view on the R-transform and related asymptotics of spherical integrals* J. Funct. Anal. 222, 2, (2005), 435–490.

[13] A. Guionnet *Large random matrices: lectures on macroscopic asymptotics* École d'Été de probabilités de Saint-Flour XXXVI–2006, Lect. Notes Math. 1957 (2009).

[14] F. Hiai, D. Petz *The semicircle law, free random variables, and entropy*, Amer. Math. Soc., Mathematical Surveys and Monographs Volume 77, 2000.
[15] R.A. Horn, C.R. Johnson *Matrix analysis* Cambridge University Press, Cambridge, 1985
[16] Y. Kabashima *Inference from correlated patterns: a unified theory for perceptron learning and linear vector channels*, Journal of Physics: Conference Series 95 (2008) 012001.
[17] M. Maïda *Large deviations for the largest eigenvalue of rank one deformation of Gaussian ensembles*, Elec. Jour. Prob. 12 (2007), 1131–1150.
[18] A. Nica, R. Speicher *Lectures on the combinatorics of free probability.* London Mathematical Society Lecture Note Series, 335. Cambridge University Press, Cambridge, 2006.
[19] B. Schlittgen, T. Wettig, *Generalizations of some integrals over the unitary group* J. Phys. A, 36, 3195 (2003).
[20] P. Zinn-Justin, J.-B. Zuber *On some integrals over the $U(N)$ unitary group and their large N limit.* J. Phys. A 36, 12 (2003) 3173–3193.
[21] A. Zvonkin *Matrix integrals and map enumeration: an accessible introduction.* Math. Comput. Modelling 26, 8–10 (1997), 281–304. Combinatorics and physics (Marseilles, 1995).

Florent Benaych-Georges, LPMA, UPMC Univ Paris 6, Case courier 188, 4, Place Jussieu, 75252 Paris Cedex 05, France, and CMAP, École Polytechnique, route de Saclay, 91128 Palaiseau Cedex, France.

E-mail address: `florent.benaych@upmc.fr`